

Generalizing Bell-type and Leggett-Garg-type Inequalities to Systems with Signaling

Ehtibar N. Dzhafarov
Purdue University

Janne V. Kujala
University of Jyväskylä

Contextuality means non-existence of a joint distribution for random variables recorded under mutually incompatible conditions, subject to certain constraints imposed on how the identity of these variables may change across these conditions. In simple quantum systems contextuality is indicated by violations of Bell-type or Leggett-Garg-type inequalities. These inequalities, however, are predicated on the assumption of no-signaling, defined as invariance of the distributions of measurement results with respect to other (e.g., earlier in time) measurements' settings. Signaling makes the inequalities inapplicable: a non-signaling system with any degree of contextuality, however high, loses any relation to this concept as soon as it exhibits any degree of signaling, however small. This is unsatisfactory. We describe a principled way of defining and measuring contextuality in arbitrary systems with random outputs, whether signaling is absent or present.

KEYWORDS: Bell/CHSH inequalities; contextuality; EPR/Bohm paradigm; Leggett-Garg inequalities; signaling.

1. INTRODUCTION

Contextuality can be defined in purely probabilistic terms, for abstract systems with random outputs recorded under different (mutually incompatible) conditions [1–7]. Consider, e.g., (X_1, Y_1, Z_1, \dots) recorded under condition c_1 , (X_2, Y_2, Z_2, \dots) recorded under condition c_2 , etc. The notion of contextuality involves a hypothesis that certain random variables preserve their identity across some of the different conditions: e.g., that $X_1 = X_2$. The system exhibits no contextuality (with respect to this hypothesis) if all the random variables (X_i, Y_i, Z_i, \dots) across different values of i can be viewed as jointly distributed with X_1 and X_2 being always equal to each other. In the Kolmogorovian probability theory, being jointly distributed is equivalent to the random outputs being (measurable) functions of one and the same (“hidden”) random variable λ [8]:

$$X_i = x_i(\lambda), Y_i = y_i(\lambda), Z_i = z_i(\lambda), \dots \quad (1)$$

The constraint $X_1 = X_2$ means

$$\Pr[X_1 \neq X_2] = \Pr[\lambda : x_1(\lambda) \neq x_2(\lambda)] = 0. \quad (2)$$

As a well-known example, in the simplest Alice-Bob EPR/Bohm paradigm [9, 10], the four mutually incompatible conditions (α_i, β_j) are formed by Alice's settings α_1 or α_2 combined with Bob's settings β_1 or β_2 . Under each condition (α_i, β_j) , Alice and Bob record spins represented by binary (± 1) random variables A_{ij} and B_{ij} , respectively. We will refer to a system with this input-output relation as a Bell-system. It involves eight random variables, with the joint distribution being known for each pair (A_{ij}, B_{ij}) but not across different pairs. The identity hypothesis here is that $A_{i1} = A_{i2}$ for $i = 1, 2$, and $B_{1j} = B_{2j}$ for $j = 1, 2$. Stated rigorously, if one can impose a joint distribution on all eight random variables

consistent with the known distributions of (A_{ij}, B_{ij}) and constrained by the requirement

$$\Pr[A_{i1} \neq A_{i2}] = \Pr[B_{1j} \neq B_{2j}] = 0, \quad i, j \in \{1, 2\}, \quad (3)$$

then the Bell system exhibits no contextuality.

Similarly, in the simplest Leggett-Garg paradigm [11], there are three mutually exclusive conditions (t_1, t_2) , (t_1, t_3) , and (t_2, t_3) , formed by three fixed time moments $t_1 < t_2 < t_3$. The two binary (± 1) random outputs jointly recorded at moments $t_i < t_j$ can be denoted Q_{ij} and Q_{ji} , respectively. We will refer to a system with this input-output relation as an LG-system. It involves six random variables, with the joint distribution known for each pair (Q_{ij}, Q_{ji}) but not across different pairs. The identity hypothesis here is that $Q_{12} = Q_{13}$, $Q_{21} = Q_{23}$, and $Q_{31} = Q_{32}$. The LG-system exhibits no contextuality if one can impose a joint distribution on all six random variables consistent with the known distributions of the pairs (Q_{ij}, Q_{ji}) and subject to

$$\Pr[Q_{12} \neq Q_{13}] = \Pr[Q_{21} \neq Q_{23}] = \Pr[Q_{31} \neq Q_{32}] = 0. \quad (4)$$

The issue we take on in this paper is related to the fact that non-contextuality defined as above implies the condition known as marginal selectivity [8, 12] or no-signaling [13, 14]: obviously, any set of random variables whose identity is preserved across different conditions preserves its distribution across these conditions. For the Bell-systems, no-signaling means, using $\langle \cdot \rangle$ for expected value,

$$\langle A_{i1} \rangle = \langle A_{i2} \rangle, \quad \langle B_{1j} \rangle = \langle B_{2j} \rangle, \quad i, j \in \{1, 2\}, \quad (5)$$

while for the LG-systems it means

$$\langle Q_{ij} \rangle = \langle Q_{ij'} \rangle, \quad i, j, j' \in \{1, 2, 3\}, \quad i \neq j, \quad i \neq j'. \quad (6)$$

The necessary and sufficient condition for non-contextuality in the two types of systems are obtained as

conjunctions of the no-signaling requirements just given with certain inequalities involving jointly distributed pairs: for the Bell-systems it is the conjunction of (5) with the CHSH inequality [15]

$$\max_{i,j \in \{1,2\}} \left| \langle A_{11}B_{11} \rangle + \langle A_{12}B_{12} \rangle + \langle A_{21}B_{21} \rangle + \langle A_{22}B_{22} \rangle - 2 \langle A_{ij}B_{ij} \rangle \right| \leq 2, \quad (7)$$

while for the LG-systems it is the conjunction of (6) with the Leggett-Garg-Suppes-Zanotti (LGSZ) inequality [11, 16, 17]

$$\begin{aligned} -1 &\leq \langle Q_{12}Q_{21} \rangle + \langle Q_{13}Q_{31} \rangle + \langle Q_{23}Q_{32} \rangle \\ &\leq 1 + 2 \min \{ \langle Q_{12}Q_{21} \rangle, \langle Q_{13}Q_{31} \rangle, \langle Q_{23}Q_{32} \rangle \}. \end{aligned} \quad (8)$$

The inequalities are logically independent of the corresponding no-signaling conditions: one can construct examples of systems with all four combinations of truth values for (5) and (7), or for (6) and (8) [18].

Logically, then, we should consider a Bell-system exhibiting contextuality if either CHSH inequalities (7) are violated or no-signaling condition (5) is violated (or both); and analogously for the LG-systems. However, to posit that any instance of signaling constitutes contextuality amounts to unreasonably expanding the meaning of contextuality, and it contradicts the common usage. If changes in Bob's setting somehow change the distribution of spins recorded by Alice under a fixed setting (assuming the two are separated by a time-like interval), the natural language to use is that of direct cross-influences rather than contextuality. But it is equally unsatisfactory to declare (non-)contextuality undefined whenever signaling is present. Consider, e.g., a Bell system with

$$\begin{aligned} \langle A_{11}B_{11} \rangle &= \langle A_{12}B_{12} \rangle = \langle A_{21}B_{21} \rangle = -\langle A_{22}B_{22} \rangle = \delta, \\ \langle A_{11} \rangle &= \langle B_{11} \rangle = \langle A_{12} \rangle = \langle B_{12} \rangle = \langle A_{21} \rangle = \langle B_{21} \rangle = 0, \\ \langle A_{22} \rangle &= -\langle B_{22} \rangle = \varepsilon. \end{aligned} \quad (9)$$

It satisfies the no-signaling condition (5) if and only if $\varepsilon = 0$. In this case, for any $\delta > 1/2$, it violates CHSH inequalities (7), indicating thereby contextuality. If the degree of contextuality is measured as proportional to the excess of the left-hand side of (7) over 2, the maximum contextuality allowed by quantum mechanics [19] is achieved at $\delta = 1/\sqrt{2}$, whereas $\delta = 1$ represents a Bell-system with maximum contextuality algebraically possible [20]. But as soon as ε differs from zero, however slightly, contextuality changes from a very high (even highest possible) level to being undefined. Among other things, this creates difficulties for statistical analysis of contextuality, where one can never establish with certainty that equalities (5) and (6) hold precisely.

In this paper we propose a new definition and new measure of contextuality that overcome this difficulty: even in the presence of direct cross-influences (say, from Bob's setting to Alice's measurements and vice versa) one can identify and compute the degree of contextual influences "on top of" the direct cross-influences.

2. CRITERION FOR (NON)CONTEXTUALITY

The main idea is this: contextuality is present if random variables recorded under different conditions cannot be presented as a single system of jointly distributed random variables, provided their identity across different conditions changes as little as it is possible in view of the observed differences between marginal distributions (i.e., in view of signaling).

For a Bell-system, we consider the vector of probabilities [21]

$$C = \left(\begin{array}{c} \Pr[A_{11} \neq A_{12}], \Pr[A_{21} \neq A_{22}], \\ \Pr[B_{11} \neq B_{21}], \Pr[B_{12} \neq B_{22}] \end{array} \right), \quad (10)$$

and find the minimum possible values of these probabilities allowed by the system's marginal expectations

$$\left(\begin{array}{c} \langle A_{11} \rangle, \langle A_{12} \rangle, \langle A_{21} \rangle, \langle A_{22} \rangle, \\ \langle B_{11} \rangle, \langle B_{21} \rangle, \langle B_{12} \rangle, \langle B_{22} \rangle \end{array} \right). \quad (11)$$

Denote this vector C by C_0 . It is specified as follows.

Lemma 1. *Given marginals (11) of a Bell-system,*

$$C_0 = \left(\begin{array}{c} \frac{1}{2} |\langle A_{11} \rangle - \langle A_{12} \rangle|, \frac{1}{2} |\langle A_{21} \rangle - \langle A_{22} \rangle|, \\ \frac{1}{2} |\langle B_{11} \rangle - \langle B_{21} \rangle|, \frac{1}{2} |\langle B_{12} \rangle - \langle B_{22} \rangle| \end{array} \right). \quad (12)$$

The proof of this and subsequent formal statements is relegated to Appendix. Note that under no-signaling we have $C_0 = \mathbf{0}$, in accordance with (3). The question we ask is whether this C_0 is compatible with the observed distributions of the pairs (A_{ij}, B_{ij}) . If it is, the Bell-system exhibits no contextuality. If it is not, then contextuality is present, and a measure of its degree is easily computed as shown below.

The compatibility of C_0 with the observed pairs of random outputs means that a joint distribution can be imposed on all eight random variables so that it is consistent with both C_0 and the observed pairs. In other words, each of the 2^8 possible combinations

$$A_{11} = \pm 1, B_{11} = \pm 1, \dots, A_{22} = \pm 1, B_{22} = \pm 1 \quad (13)$$

can be assigned a probability, so that the probabilities for all combinations containing, say, $A_{12} = 1$ and $B_{12} = -1$ sum to the observed $\Pr[A_{12} = 1, B_{12} = -1]$; and the probabilities for all combinations containing unequal values of, say, B_{12} and B_{22} sum to $\Pr[B_{12} \neq B_{22}]$ in C_0 .

Theorem 2 (non-contextuality criterion for Bell-systems). *A Bell-system exhibits no contextuality, i.e., C_0 in (12) is compatible with the observed pairs $(A_{ij}, B_{ij})_{i,j \in \{1,2\}}$, if and only if*

$$\max_{i,j \in \{1,2\}} \left| \langle A_{11}B_{11} \rangle + \langle A_{12}B_{12} \rangle + \langle A_{21}B_{21} \rangle + \langle A_{22}B_{22} \rangle - 2 \langle A_{ij}B_{ij} \rangle \right| \leq 2(1 + \Delta_0), \quad (14)$$

where Δ_0 is the sum of the components of C_0 ,

$$\Delta_0 = \frac{1}{2} \left(|\langle A_{11} \rangle - \langle A_{12} \rangle| + |\langle A_{21} \rangle - \langle A_{22} \rangle| + |\langle B_{11} \rangle - \langle B_{21} \rangle| + |\langle B_{12} \rangle - \langle B_{22} \rangle| \right). \quad (15)$$

For the LG-system the situation is analogous. We consider a vector of probabilities

$$C' = (\Pr[Q_{12} \neq Q_{13}], \Pr[Q_{21} \neq Q_{23}], \Pr[Q_{31} \neq Q_{32}]) \quad (16)$$

and determine C'_0 with the minimum values of these probabilities allowed by the system's marginals

$$(\langle Q_{12} \rangle, \langle Q_{13} \rangle, \langle Q_{21} \rangle, \langle Q_{23} \rangle, \langle Q_{31} \rangle, \langle Q_{32} \rangle). \quad (17)$$

Lemma 3. *Given marginals (17) of an LG-system,*

$$C'_0 = \left(0, \frac{1}{2} |\langle Q_{21} \rangle - \langle Q_{23} \rangle|, \frac{1}{2} |\langle Q_{31} \rangle - \langle Q_{32} \rangle|\right). \quad (18)$$

Note that, by causality considerations, $|\langle Q_{12} \rangle - \langle Q_{13} \rangle|$ in C'_0 must equal zero (but it need not be in a generalized treatment, if t_1, t_2, t_3 are treated as labels other than time moments).

Theorem 4 (non-contextuality criterion for LG-systems). *An LG-system exhibits no contextuality, i.e., C'_0 in (18) is compatible with the observed pairs $(Q_{12}, Q_{21}), (Q_{13}, Q_{31}), (Q_{23}, Q_{32})$, if and only if*

$$\begin{aligned} -1 - 2\Delta'_0 &\leq \langle Q_{12}Q_{21} \rangle + \langle Q_{13}Q_{31} \rangle + \langle Q_{23}Q_{32} \rangle \\ &\leq 1 + 2\Delta'_0 + 2 \max \{ \langle Q_{12}Q_{21} \rangle, \langle Q_{13}Q_{31} \rangle, \langle Q_{23}Q_{32} \rangle \}, \end{aligned} \quad (19)$$

where Δ'_0 is the sum of the components of C'_0 ,

$$\Delta'_0 = \frac{1}{2} (|\langle Q_{21} \rangle - \langle Q_{23} \rangle| + |\langle Q_{31} \rangle - \langle Q_{32} \rangle|). \quad (20)$$

Under no-signaling condition, Δ_0 and Δ'_0 are zero, and Theorems 2 and 4 reduce to the traditional non-contextuality criteria (5)-(7) and (6)-(8), respectively. Note also that a Bell-system with $\Delta_0 > 1$ and an LG-system with $\Delta'_0 > 1$ are necessarily non-contextual, as (14) and, respectively, (19) then cannot be violated.

3. DEGREE OF CONTEXTUALITY UNDER SIGNALING

A measure of contextuality is based on the same compatibility-under-constraints considerations as the criteria just derived. For a Bell-system, let Δ_{\min} be the minimum value of

$$\Delta = \frac{\Pr[A_{11} \neq A_{12}] + \Pr[A_{21} \neq A_{22}]}{\Pr[B_{11} \neq B_{21}] + \Pr[B_{12} \neq B_{22}]} \quad (21)$$

that is compatible with the observed pairs $(A_{ij}, B_{ij})_{i,j \in \{1,2\}}$. It follows from the previous that the system exhibits contextuality if and only if this Δ_{\min} exceeds the value of Δ_0 in (15). It is natural therefore to define the degree of contextuality in a Bell system as

$$\max(0, \Delta_{\min} - \Delta_0) \quad (22)$$

This value is well-defined and given by

Theorem 5 (contextuality degree in Bell-systems). *The degree of contextuality in a Bell-system is*

$$\max \left\{ 0, \frac{1}{2} \max_{i,j \in \{1,2\}} \left| \frac{\langle A_{11}B_{11} \rangle + \langle A_{12}B_{12} \rangle}{\langle A_{21}B_{21} \rangle + \langle A_{22}B_{22} \rangle - 2 \langle A_{ij}B_{ij} \rangle} \right| - 1 - \Delta_0 \right\}. \quad (23)$$

The degree of contextuality thus is always nonnegative. It equals zero if and only if $\Delta_{\min} = \Delta_0$, which is equivalent to (14). Returning to our motivating example (9), the degree of contextuality there is $\max(0, 2\delta - 1 - 2|\varepsilon|)$, changing continuously with ε .

For LG-systems the degree of contextuality is defined analogously, as

$$\max(0, \Delta'_{\min} - \Delta'_0),$$

where Δ'_{\min} is the smallest value of

$$\Delta' = \Pr[Q_{12} \neq Q_{13}] + \Pr[Q_{23} \neq Q_{21}] + \Pr[A_{32} \neq A_{31}] \quad (24)$$

compatible with the observed pairs $(Q_{ij}, Q_{ji})_{i < j \in \{1,2,3\}}$.

Theorem 6 (contextuality degree in LG-systems). *The degree of contextuality in an LG-system is*

$$\max \left\{ 0, \frac{1}{2} \max \left\{ \begin{array}{l} \pm \langle Q_{12}Q_{21} \rangle \pm \langle Q_{13}Q_{31} \rangle \\ \pm \langle Q_{23}Q_{32} \rangle : \\ \text{number of minuses} \\ \text{is odd} \end{array} \right\} - \frac{1}{2} - \Delta'_0 \right\}. \quad (25)$$

Appendix: Proofs

We use the convenient notion of a (probabilistic) connection [5, 22], as defined in Fig. 1. We also make use of two functions: for any natural r , $s_0(x_1, \dots, x_{2r})$ stands for $\max \{(\pm x_1 \dots \pm x_{2r}) : \# \text{ of minuses is even}\}$, and $s_1(x_1, \dots, x_r)$ denotes $\max \{(\pm x_1 \dots \pm x_r) : \# \text{ of minuses is odd}\}$.

Proof of Lemma 1. Consider, e.g., the distribution of the connection (A_{11}, A_{12}) :

| | $A_{12} = +1$ | $A_{12} = -1$ | |
|---------------|-----------------------|-----------------------|-------|
| $A_{11} = +1$ | p | $\Pr[A_{11} = 1] - p$ | (A.1) |
| $A_{11} = -1$ | $\Pr[A_{12} = 1] - p$ | \dots | |

The largest possible value for p is $\min \{\Pr[A_{11} = 1], \Pr[A_{12} = 1]\}$, whence the minimum of $\Pr[A_{11} \neq A_{12}]$, which is the sum of the entries on the minor diagonal, is $|\Pr[A_{11} = 1] - \Pr[A_{12} = 1]| = \frac{1}{2} |\langle A_{11} \rangle - \langle A_{12} \rangle|$. \square

Lemma 3 is proved in the same way.

The theorems of this paper are based on the following four lemmas. Their proofs are computer-assisted, as they boil down to symbolically solving large systems of linear inequalities.

Lemma A.1. *The necessary and sufficient condition for the connections $((A_{i1}, A_{i2}), (B_{1j}, B_{2j}))_{i,j \in \{1,2\}}$ to be compatible with the observed pairs $(A_{ij}, B_{ij})_{i,j \in \{1,2\}}$ is*

$$\begin{aligned} & s_0 (\langle A_{11} B_{11} \rangle, \langle A_{12} B_{12} \rangle, \langle A_{21} B_{21} \rangle, \langle A_{22} B_{22} \rangle) \\ & + s_1 (\langle A_{11} A_{12} \rangle, \langle B_{11} B_{21} \rangle, \langle A_{21} A_{22} \rangle, \langle B_{12} B_{22} \rangle) \leq 6, \\ & s_1 (\langle A_{11} B_{11} \rangle, \langle A_{12} B_{12} \rangle, \langle A_{21} B_{21} \rangle, \langle A_{22} B_{22} \rangle) \\ & + s_0 (\langle A_{11} A_{12} \rangle, \langle B_{11} B_{21} \rangle, \langle A_{21} A_{22} \rangle, \langle B_{12} B_{22} \rangle) \leq 6. \end{aligned} \quad (\text{A.2})$$

Proof. The joint distribution of the eight random variables A_{ij}, B_{ij} , $i, j \in \{1,2\}$, is fully described by the vector $\mathbf{q} \in [0,1]^n$, $q_1 + \dots + q_n = 1$, consisting of the probabilities of the $n = 2^8$ different combinations of the values of the 8 random variables. We define a vector $\mathbf{p} \in [0,1]^m$, $m = 32$, consisting of the 16 observed probabilities $\Pr[A_{ij} = a, B_{ij} = b]$ and the 16 connection probabilities $\Pr[A_{i1} = a, A_{i2} = a']$ and $\Pr[B_{1j} = b, B_{2j} = b']$, where $a, a', b, b' \in \{-1,1\}$ and $i, j \in \{1,2\}$. The observed probabilities are compatible with the connection probabilities if and only if there exists an n -vector $\mathbf{q} \geq 0$ (componentwise) such that $\mathbf{p} = M\mathbf{q}$, where $M \in \{0,1\}^{m \times n}$ determines which components of \mathbf{q} sum to each component of \mathbf{p} . As described in Text S3 of Ref. [23], the set of vectors \mathbf{p} forms a polytope whose vertices are given by the columns of M and whose half-space representation can be obtained by a facet enumeration algorithm. This half-space representation consists of 160 inequalities, as well as 16 equations ensuring that the marginals of the observed probabilities agree with those of the connections and the probabilities are properly normalized. Expressing the probabilities in \mathbf{p} in terms of the observed and connection expectations $(\langle A_{ij} B_{ij} \rangle, \langle A_{ij} \rangle, \langle B_{ij} \rangle, \langle A_{i1} A_{i2} \rangle, \langle B_{1j} B_{2j} \rangle)$, $i, j \in \{1,2\}$, the 16 equations become identically true (the parameterization alone guarantees them), and of the 160 inequalities, 128 turn into exactly those represented by (A.2); the remaining 32 inequalities need not be listed as they are constraints of the form $-1 + |\langle A \rangle + \langle B \rangle| \leq \langle AB \rangle \leq 1 - |\langle A \rangle - \langle B \rangle|$, trivially following from the non-negativity of probabilities. \square

This proof is different from the similar result in Ref. [23] in that the parameterization for the probabilities in \mathbf{p} is more general (allowing for arbitrary marginals of the eight random variables) and so we obtain a more general condition for the compatibility of observed and connection probabilities.

Lemma A.2. *The necessary and sufficient condition for the connections $(Q_{12}, Q_{13}), (Q_{21}, Q_{23}), (Q_{31}, Q_{32})$ to be compatible with the observed pairs $(Q_{12}, Q_{21}), (Q_{13}, Q_{31}), (Q_{23}, Q_{32})$ is*

$$s_1 \left(\langle Q_{12} Q_{21} \rangle, \langle Q_{13} Q_{31} \rangle, \langle Q_{23} Q_{32} \rangle, \langle Q_{12} Q_{13} \rangle, \langle Q_{21} Q_{23} \rangle, \langle Q_{31} Q_{32} \rangle \right) \leq 4. \quad (\text{A.3})$$

The proof is analogous to that of Lemma A.1.

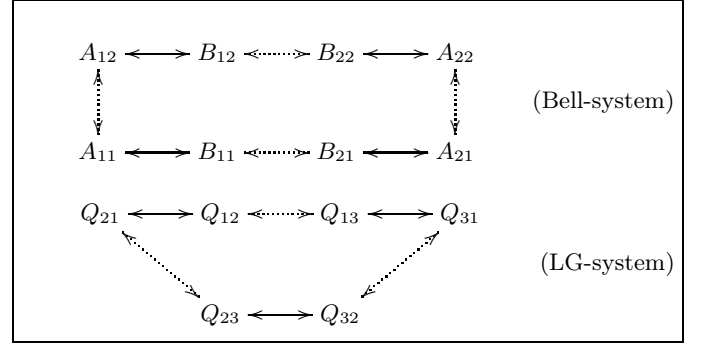


Figure A.1: Random variables involved in the Bell-system and LG-system. The pairs of random variables whose joint distributions are empirically observed, e.g., (A_{12}, B_{12}) and (Q_{12}, Q_{21}) , are indicated by solid double-arrows. The pairs of random variables forming probabilistic connections (with unobservable joint distributions) are indicated by point double-arrows, e.g., (A_{11}, A_{12}) and (Q_{12}, Q_{13}) . Lemmas 1 and 3 are about connections whose components are as close to being identical as possible; Theorems 2 and 4 are about connections compatible with the observed pairs.

Lemma A.3. *If the connections $((A_{i1}, A_{i2}), (B_{1j}, B_{2j}))_{i,j \in \{1,2\}}$ are compatible with the observed pairs $(A_{ij}, B_{ij})_{i,j \in \{1,2\}}$, then, with Δ defined as in (21),*

$$\begin{aligned} \Delta & \geq -1 + \frac{1}{2} s_1 \left(\langle A_{11} B_{11} \rangle, \langle A_{12} B_{12} \rangle, \langle A_{21} B_{21} \rangle, \langle A_{22} B_{22} \rangle \right), \\ \Delta & \geq \frac{1}{2} \left(|\langle A_{11} \rangle - \langle A_{12} \rangle| + |\langle A_{21} \rangle - \langle A_{22} \rangle| \right. \\ & \quad \left. + |\langle B_{11} \rangle - \langle B_{21} \rangle| + |\langle B_{12} \rangle - \langle B_{22} \rangle| \right), \\ \Delta & \leq 5 - \frac{1}{2} s_1 \left(\langle A_{11} B_{11} \rangle, \langle A_{12} B_{12} \rangle, \langle A_{21} B_{21} \rangle, \langle A_{22} B_{22} \rangle \right), \\ \Delta & \leq 4 - \frac{1}{2} \left(|\langle A_{11} \rangle + \langle A_{12} \rangle| + |\langle A_{21} \rangle + \langle A_{22} \rangle| \right. \\ & \quad \left. + |\langle B_{11} \rangle + \langle B_{21} \rangle| + |\langle B_{12} \rangle + \langle B_{22} \rangle| \right). \end{aligned} \quad (\text{A.4})$$

Conversely, if these inequalities are satisfied for a given value of Δ , then the connection distributions can always be chosen so that yield this value of Δ and are compatible with the distributions of the observed pairs.

Proof. Given the 160 inequalities of Lemma A.1 (characterizing the compatibility of the connections with the observed pairs), we add to this linear system the equation defining Δ in terms of the expectations $(\langle A_{i1} A_{i2} \rangle, \langle B_{1j} B_{2j} \rangle, \langle A_{ij} \rangle, \langle B_{ij} \rangle)_{i,j \in \{1,2\}}$. Then we use this equation to eliminate one of the connection expectation variables $(\langle A_{i1} A_{i2} \rangle, \langle B_{1j} B_{2j} \rangle)_{i,j \in \{1,2\}}$ from the system (by solving the equation for this variable and then substituting the solution everywhere else). After that, we eliminate the three remaining connection expectation variables one by one using the Fourier-Motzkin elimination algorithm [24]. Then we remove any redundant inequalities from the system by linear programming using

the algorithm described in Ref. [23], Text S3. After having eliminated all connection expectation variables and having deleted the inequalities following from the non-negativity of probabilities, we are left with the system (A.4). The Fourier-Motzkin elimination algorithm guarantees that the resulting system has a solution precisely when the original system has a solution with some values of the eliminated variables. \square

Lemma A.4. *If the connections $(Q_{12}, Q_{13}), (Q_{21}, Q_{23}), (Q_{31}, Q_{32})$ are compatible with the observed pairs $(Q_{12}, Q_{21}), (Q_{13}, Q_{31}), (Q_{23}, Q_{32})$, then, with Δ' defined as in ((24)),*

$$\begin{aligned} \Delta' &\geq -\frac{1}{2} + \frac{1}{2}s_1(\langle Q_{12}Q_{21} \rangle, \langle Q_{13}Q_{31} \rangle, \langle Q_{23}Q_{32} \rangle), \\ \Delta' &\geq \frac{1}{2} \left(|\langle Q_{12} \rangle - \langle Q_{13} \rangle| + |\langle Q_{21} \rangle - \langle Q_{23} \rangle| + |\langle Q_{31} \rangle - \langle Q_{32} \rangle| \right), \\ \Delta' &\leq \frac{7}{2} - \frac{1}{2}s_1(\langle Q_{12}Q_{21} \rangle, \langle Q_{13}Q_{31} \rangle, \langle Q_{23}Q_{32} \rangle), \\ \Delta' &\leq 3 - \frac{1}{2} \left(|\langle Q_{12} \rangle + \langle Q_{13} \rangle| + |\langle Q_{21} \rangle + \langle Q_{23} \rangle| + |\langle Q_{31} \rangle + \langle Q_{32} \rangle| \right). \end{aligned} \quad (\text{A.5})$$

Conversely, if these inequalities are satisfied for a given value of Δ' , then the connection distributions can always be chosen so that yield this value of Δ' and are compatible with the distributions of the observed pairs.

The proof is analogous to that of Lemma (A.3).

Proof of Theorems 2 and 5. Inequalities (A.4) in Lemma A.2 can be easily checked to be mutually compatible, whence Δ_{\min} is the larger of the two right-hand expressions in the first and third of them. Note that $s_1(\dots)$ is the same as $\max|\dots|$ -part of (23). This proves Theorem 2, and Theorem 5 follows as an explication of $\Delta_{\min} = \Delta_0$. \square

The proofs of Theorems 4 and 6 follows from Lemma A.4 analogously.

This work was supported by NSF grant SES-1155956. The authors are grateful to J. Acacio de Barros, Gary Oas, Jan-Åke Larsson, and Guido Bacciagaluppi for helpful discussions of issues related to probabilistic contextuality.

-
- [1] E. Specker (1960). Die Logik Nicht Gleichzeitig Entscheiderbarer Aussagen. *Dialectica* 14: 239–246 (English translation by M.P. Seevinck available as arXiv:1103.4537.)
 - [2] J.-Å. Larsson (2022). A Kochen-Specker inequality. *Europhys. Lett.*, 58: 799–805.
 - [3] A.Yu. Khrennikov (2008). EPR–Bohm experiment and Bell’s inequality: Quantum physics meets probability theory. *Theor. Math. Phys.* 157: 1448–1460.
 - [4] A. Cabello (2013). Simple explanation of the quantum violation of a fundamental inequality. *Phys. Rev. Lett.* 110: 060402.
 - [5] E.N. Dzhafarov & J.V. Kujala (2014). A qualified Kolmogorovian account of probabilistic contextuality. *Lecture Notes in Computer Science* 8369: 201–212.
 - [6] E.N. Dzhafarov & J.V. Kujala (2014). Embedding quantum into classical: contextualization vs conditionalization. *PLoS One* 9(3): e92818. doi:10.1371/journal.pone.0092818.
 - [7] E.N. Dzhafarov & J.V. Kujala (in press). Contextuality is about identity of random variables. *Phys. Scripta* (available as arXiv:1405.2116).
 - [8] E.N. Dzhafarov & J.V. Kujala (2013). Probability, random variables, and selectivity. arXiv:1312.2239.
 - [9] J.F. Clauser, M.A. Horne, A. Shimony, & R.A. Holt (1969). Proposed experiment to test local hidden-variable theories. *Phys. Rev. Lett.* 23: 880–884.
 - [10] J.F. Clauser & M.A. Horne (1974). Experimental consequences of objective local theories. *Phys. Rev. D* 10: 526–535.
 - [11] A.J. Leggett & A. Garg (1985). Quantum mechanics versus macroscopic realism: Is the flux there when nobody looks? *Phys. Rev. Lett.* 54: 857–860.
 - [12] E.N. Dzhafarov & J.V. Kujala (2013). Order-distance and other metric-like functions on jointly distributed random variables. *Proc. Amer. Math. Soc.* 141(9): 3291–3301.
 - [13] J. Cereceda (2000). Quantum mechanical probabilities and general probabilistic constraints for Einstein–Podolsky–Rosen–Bohm experiments. *Found. Phys. Lett.* 13: 427–442.
 - [14] Ll. Masanes, A. Acin, & N. Gisin (2006). General properties of nonsignaling theories. *Phys. Rev. A* 73: 012112.
 - [15] A. Fine (1982). Hidden variables, joint probability, and the Bell inequalities. *Phys. Rev. Lett.* 48: 291–295.
 - [16] P. Suppes & M. Zanotti (1981). When are probabilistic explanations possible? *Synthese* 48: 191–199.
 - [17] Suppes & Zanotti [16] prove (8) for uniform marginals only, $\langle Q_{ij} \rangle = 0$, while Leggett & Garg [11] use only one part of the double-inequality. The generality of (8) follows as a special case from our Theorem 4.
 - [18] E.N. Dzhafarov & J.V. Kujala (2014). On selective influences, marginal selectivity, and Bell/CHSH inequalities. *Topics Cog. Sci.* 6: 121–128.
 - [19] B.S. Tsirelson (1980). Quantum generalizations of Bell’s inequality. *Lett. Math. Phys.* 4: 93–100.
 - [20] S. Popescu & D. Rohrlich (1994). Quantum nonlocality as an axiom. *Found. Phys.* 24: 379–385.
 - [21] The idea of using such probabilities to define contextuality was first introduced in a work by Larsson [2], in relation to Kochen-Specker inequalities. Within the framework of our contextuality-by-default approach [5–7, 22], these probabilities characterize probabilistic connections (see Appendix).
 - [22] E.N. Dzhafarov & J.V. Kujala (2014). No-Forcing and No-Matching theorems for classical probability applied to quantum mechanics. *Found. Phys.* 44: 248–265.
 - [23] E.N. Dzhafarov & J.V. Kujala (2013). All-possible-couplings approach to measuring probabilistic context. *PLoS ONE* 8(5): e61712. doi:10.1371/journal.pone.0061712.

- [24] G.B. Dantzig, B.C. Eaves (1972). Fourier-Motzkin elimination and its dual. J. Comb. Theory A 14: 288–297.